 DOI: 10.1111/risa.12036

An Approach for Optimal Allocation of Safety Resources: Using the Knapsack Problem to Take Aggregated Cost-Efficient Preventive Measures

Genserik L. L. Reniers[1,2,*] and Kenneth Sörensen[1]

On the basis of the combination of the well-known knapsack problem and a widely used risk management technique in organizations (that is, the risk matrix), an approach was developed to carry out a cost-benefits analysis to efficiently take prevention investment decisions. Using the knapsack problem as a model and combining it with a well-known technique to solve this problem, bundles of prevention measures are prioritized based on their costs and benefits within a predefined prevention budget. Those bundles showing the highest efficiencies, and within a given budget, are identified from a wide variety of possible alternatives. Hence, the approach allows for an optimal allocation of safety resources, does not require any highly specialized information, and can therefore easily be applied by any organization using the risk matrix as a risk ranking tool.

KEY WORDS: Cost-benefits analysis; knapsack problem; prevention measures; risk matrix

1. INTRODUCTION

Safety risks are present throughout organizations, going hand in hand with all daily activities. Such risks may lead to undesired events and to injuries, detriment, and damage. To deal with these risks, organizations use risk management methods. Since there are more risks in any organization than one could handle at one time, risk management implies the task of setting priorities. It is imperative for companies to be able to reduce and control the wide variety of existing safety risks in a cost-efficient way. Risks are reduced through risk reduction policies and prevention measures. To determine an optimal set of prevention measures, organizations need to take into account in a systematic way the measures' costs as well as their (hypothetical) benefits. Prevention benefits can be calculated by determining the difference between hypothetical accident costs before and after implementing precaution measures. In other words, the *avoided accident costs* are calculated, or the financial benefits of *accidents not happening*. The approach builds on these insights.

Three types of accidents can be discerned: accidents where a lot of historical data is available (type I), accidents where little or extremely little historical data is available (type II), and accidents where no historical data is available (type III). Whereas type I accidents imply most work-related accidents such as falling, small fires, slipping, etc., type II accidents can be regarded as catastrophes with major consequences and often with multiple fatalities. Type II accidents do occur on a (semi-)regular basis in a worldwide perspective, and large fires, large releases, explosions, toxic clouds, etc. belong to this class of accidents. Type III accidents are often "true major disasters" in terms of the losses of lives and/or in terms of economic devastation. These accidents often become part of the collective memory of humankind.

[1]Faculty of Applied Economic Sciences, Research Groups ARGoSS and ANT/OR, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium.

[2]Centre for Economics and Corporate Sustainability (CEDON), HUB, KULeuven, Warmoesberg 26, 1000 Brussels, Belgium.

*Address correspondence to Genserik L. L. Reniers, Centre for Economics and Corporate Sustainability (CEDON), HUB, KULeuven, Warmoesberg 26, 1000 Brussels, Belgium; genserik.reniers@kuleuven.be.

To prevent type I accidents, risk management techniques and practices are widely available. The interested reader is, for example, referred to Greenberg and Cramer[1] and CCPS.[2] Statistical analyses based on past accidents can be drafted to predict possible future type I accidents, indicating the prevention measures that need to be taken to prevent such accidents.

Type II accidents are much more difficult to predict. They cannot be predicted via standard statistical analyses since the frequency with which these events happen is too low. Nevertheless, some information is available, but advanced statistical methods should be employed with great care and healthy criticism, making it difficult to investigate and analyze this information. Hence, managing such risks is based on the scarce data that are available and on extrapolations, assumptions, and expert opinions.[3] The third type of accidents is simply impossible to predict. No information is available about them and they only happen extremely rarely. Such accidents can also be called "black swan accidents."[4]

In this article, only type I risks are envisioned; hence our approach is developed for nonmajor accident risks with historical data available. There should be an adequate amount of data available so as to be able to use the risk matrix as a risk management tool. Our objective is then to elaborate a decision support approach, providing its user with an idea of how to assess a safety budget in an optimal way for different bundles of precaution measures.

Section 2 defines the risk matrix and discusses its construct. Cost-benefits analysis as an instrument to determine (bundles of) precaution measures is also explained. The knapsack-based approach is developed and discussed in Section 3, and an illustrative example is given. The last section contains the conclusions of this article.

2. BUILDING BLOCKS FOR THE APPROACH

2.1. Risk Matrix

To describe and to analyze all possible accident scenarios in an organization, and to take precaution measures for every potential incident is, given the limitations in time and resources of organizations, regretfully not practicable and not achievable. Therefore, an adequate ranking of possible incidents is desired, before carrying out more thorough risk assessments and/or prevention measures. Cox[5] indicates that risk scores or so-called risk indices resulting from using simple input components such as "exposure, probability, consequences" in occupational health and safety risk management applications are widely employed in industrial practice to this end. Such risk indices determine the relative importance of different risks that an organization faces. Cox[5] discusses three such indices, that is, *risk*, *risk reduction*, and *risk reduction per cost ratio*. *Risk* is the most basic index and corresponds to the expression of a likelihood of an unwanted event, which is multiplied by the consequence of the event. *Risk reduction* is simply the product of the risk and the fraction of risk eliminated, if addressed. *Risk reduction per cost ratio* takes risk reduction and divides it by the cost needed to achieve it. Solving a combinatorial optimization problem by using a knapsack algorithm, Cox[5] shows that the risk index may achieve only a fraction (60%) of the risk reduction benefits of the risk reduction per cost ratio for the same cost. However, to be able to use the more sophisticated indices, information such as the "fraction of eliminated risk" is required, as well as the costs of prevention of every individual identified risk. If thousands and more risks have been identified in an organization, which is easily the case in industrial practice, determining this kind of specific information for every risk is very difficult to obtain and requires a substantial amount of time (thus also financial resources). Therefore, as Cox[5] rightfully mentions, most risk management software products only calculate the risk index, and they do not determine any more sophisticated indices. They display risk index outputs as so-called risk matrices, with frequency and severity categories for rows and columns.

Indeed, a concept well-known in organizations throughout the world for ranking risks is the so-called risk assessment decision matrix or in short the "risk matrix," allowing its user to make a classification of risks in a systematic and transparent way. This tool can be employed to measure and categorize risks on an informed judgment basis as to both likelihood and consequence and as to relative importance. Various definitions of the risk matrix have been proposed. Table I provides a nonexhaustive overview of definitions.

Using the risk matrix in an organization does not require any specific expertise in quantitative risk analysis methods or in data analysis. The method is therefore not expensive in its application.[10] Despite the fact that the risk matrix merely allows to roughly assess risks, many decisionmakers and consultants are convinced that the method is very useful

Table I. Nonexhaustive List of Risk Matrix Definitions

Author(s)	Risk Matrix Definition
Cook[6]	A risk matrix, in the context of safety management, is a technique for assigning a risk class to a potential accident or to a hazard in accordance with the predicted severity and absolute likelihood of the potential accident or hazard. This technique involves the construction and application of a two-dimensional matrix.
Wilkinson and David[7]	Risk matrices are tools for comparing risks relative to one another (e.g., within a single system) and hence being able to "rank" them relative to each other for the purposes of risk mitigation and the allocation of safety resources. Risk matrices are not tools for determining the tolerability, or otherwise, of individual or "single risks."
Cox, Jr.[8]	A risk matrix is a table that has several categories of "probability," "likelihood," or "frequency" for its rows (or columns) and several categories of "severity," "impact," or "consequences" for its columns (or rows, respectively). It associates a recommended level of risk, urgency, priority, or management action with each row-column pair, that is, with each cell.
Smith *et al.*[9]	Risk matrices used in the industry characterize particular risks in terms of the likelihood of occurrence, and the consequence of the actualized risk.

to make a qualitative distinction between most and least prioritized risks. The qualitative difference between risks that is based on the risk matrix is preferable over ad random decision taking.[8] Nonetheless, the risk matrix suffers from some limitations. It is not necessarily true that risk matrices provide qualitative useful information to formulate risk priorities and to identify all risks that have to be tackled and those that can be disregarded.[8] The risk matrix should be employed with great caution to investigate risks of type II and type III. This is actually the reason that the approach that is worked out in this article is recommended only to be used for type I (or nonmajor) risks.

Nonmajor risks encompass a variety of possible scenarios, leading to a number of potential consequences, either of human nature (e.g., different types of injuries, fatalities, etc.), or material devastation (e.g., production losses, installation damages, etc.). All consequences may be expressed in financial terms. Table II presents an example of such a risk matrix.

Once the hazards have been identified, the question of assigning severity and probability ratings must be addressed. A common, basic example of assigned ratings by a team on a generalized basis can be found in Table III.

The probability level F, "impossible," makes it possible to assess residual risks for cases in which the hazard is designed out of the system.

The rankings provide a quick and simple priority sorting method. The rankings are then given definitions that include, for example, definitions and recommended actions similar to those in Table IV.

It is very important that frequency estimates and consequence estimates are very well considered and carried out by experienced risk managers.

On the basis of the risk matrix explained and displayed before, a risk matrix that is divided into four consequence grades and five likelihood grades can, for example, be used. Consequence grades can be expressed in financial terms, while likelihood grades are expressed in the number of times per year that a risk leads to an accident in an organization (e.g., based on generic historical information). Financial losses linked to accident consequences may be estimated through direct and indirect costs that might occur due to the risk.[12,13] Readers interested in risk matrix configurations are referred to Greenberg and Cramer,[1] CCPS,[2] Moses and Malone,[14] Cook,[6] Levner *et al.*,[15] and Smith *et al.*[9]

Fig. 1 illustrates the risk matrix used in the remain of this article. Every cell of the risk matrix corresponds to a risk class. Per cell (thus per risk class), the financial consequence value is multiplied by the likelihood value, and the total yearly costs per risk class are determined and shown.

The numbers shown in Fig. 1 are illustrative in a sense that every company may develop its own risk matrix with its preferred numbers, applicable to the company. As a recommendation, to guarantee the usefulness of the matrix for our method designed for type I risks, we advise to use a cutoff consequence class with a maximum financial impact of €2,500,000. We use Fig. 1 in the remainder of the article to explain our approach. We assume that every type I risk can be subdivided into one of the risk classes displayed in Fig. 1. Cox[8] indicates that a certain risk in each of the cells of any risk matrix is not equally large (or small) due to the classification into risk classes. Hence, a risk cell may contain different varieties of risks. This does not pose a problem in our study, since we aim at comparing groups or bundles of risks, not individual risks.

Table II. The Risk Matrix (an Example)

Severity of consequences	Probability of Hazard					
	F Impossible	E Improbable	D Remote	C Occasional	B Probable	A Frequent
I Catastrophic					1.	
II Critical				3.	2.	
III Marginal			4.			
IV Negligible						

Risk Code/ Actions	1.	Un-acceptable	2.	Un-desirable	3.	Acceptable with controls	4.	Acceptable

Source: Based on Department of Defense.[11]

Table III. Criticality and Frequency Rating for the Risk Matrix

Severity of Consequences—Ratings		
Category:	Descriptive Word:	Results in Either:
I	Catastrophe	An onsite or an offsite death Damage and production loss greater than €750,000
II	Critical	Multiple injuries Damage and production loss between €75.000 and €750,000
III	Marginal	A single injury Damage and production loss between €7,500 and €75,000
IV	Negligible	No injuries Damage and production loss less than €7,500
Hazard Probability—Ratings		
Level:	Descriptive Word:	Definition:
A.	Frequent	Occurs more than once per year
B.	Probable	Occurs between 1 and 10 years
C.	Occasional	Occurs between 10 and 100 years
D.	Remote	Occurs between 100 and 10,000 years
E.	Improbable	Occurs less often than once per 10,000 years
F.	Impossible	Physically impossible to occur

Source: Based on Department of Defence.[11]

A discretization of the risk matrix into n cells is illustrated in Fig. 2. Every risk cell is numbered from 1 to n (in our example, $n = 20$).

The discretization of the risk matrix is carried out to have a better understanding of the approach we elaborate. The numbering will thus be used in the remainder of the article to explain our approach. The risk matrix can be refined by relating the risk classes (thus the risk cells) to a cost-benefits analysis. This way, a decision support instrument can be developed that can be used to determine, taking a certain safety budget into account, the risk reduction measures or precaution measures leading to the most optimal and cost-efficient result within an organization.

2.2. Cost-Benefits Analysis

An accident can be linked to all kinds of direct and indirect costs. It can thus be stated that by implementing a sound safety policy and by adequately taking preventative safety measures, costs can be avoided, namely, the costs of accidents that have never occurred. In this way undesired consequences can be avoided not only at the socioeconomic level but also at the environmental and the local levels. In reality, companies place little or no importance on rigorously calculating hypothetical benefits due to the complexity of the concept.[16] Table V reveals potential social costs that might accompany accidents.

Given that nonquantifiable costs are highly dependent on nongeneric data such as individuals' characteristics, company culture, and/or the company as a whole, these costs of an accident cannot be detailed in a general approach. Rather, the

Table IV. Definitions and Recommended Actions for Rankings

Ranking	Description	Required Action
1.	Unacceptable	Should be mitigated with technical measures or management procedures to a risk ranking of three or less within a specified time period such as, for example, six months.
2.	Undesirable	Should be mitigated with technical measures or management procedures to a risk ranking of three or less within a specified time period such as, for example, 12 months.
3.	Acceptable with controls	Should be verified that procedures or measures are in place.
4.	Acceptable	No mitigation action required.

Likelihood [year^{-1}]	**Cell assignments (in €/year)**			
> 1	7,500	75,000	750,000	2,500,000
> 10^{-1}	750	7,500	75,000	250,000
> 10^{-2}	75	750	7,500	25,000
> 10^{-3}	7.5	75	750	2,500
> 10^{-4}	0.75	7.5	75	250
Consequence classes / financial impact [€] →	< 7,500	< 75,000	< 750,000	< 2,500,000

Fig. 1. Risk matrix used in this article.

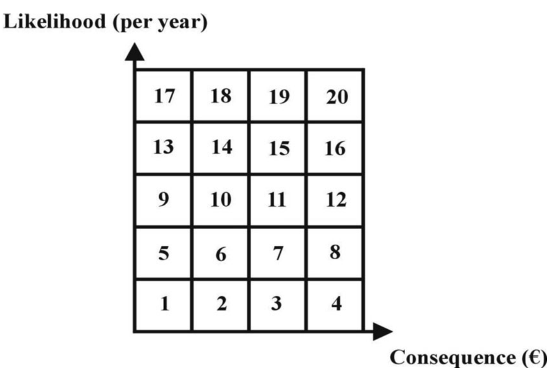

Fig. 2. Discretization of the risk matrix.

costs assert themselves when the actual costs supersede the quantifiable costs. In economics (e.g., in environment-related predicaments) monetary evaluation techniques are often used to specify nonquantifiable costs, among them the contingent valuation method and the conjoint analysis or hedonic methods.[18] In the case of nonquantifiable accident costs various studies demonstrate that these form a multiple of the quantifiable costs.[19–21]

The quantifiable socioeconomic accident costs (see Table V) can be divided into direct and indirect costs. In a situation where no accidents occur, the direct costs result in direct hypothetical benefits[3] while the indirect costs result in indirect hypothetical benefits. Resulting indirect hypothetical benefits comprise, for example, *not* having sick leave or absence from work, *not* having staff reductions, *not* experiencing labor inefficiency, and *not* experiencing change in the working environment.[22]

Selecting preventive measures can be inefficient and time consuming. Caputo *et al.*,[23] for example, present an advanced computer method to this end, solving the related combinatorial optimization problem employing a genetic algorithm. Such a method, powerful as it may be, relies on detailed knowledge of the characteristics of hazards and candidate preventive measures. Moreover, it requires the development of stochastic optimization software, which might be infeasible for the average risk analyst.

In short, reducing financial losses via precaution investments is a rather theoretical concept for an organization, due to the vast amount of specialized financial information needed. Moreover, as already stated, exact quantification of hypothetical benefits is difficult. For these reasons, organizations often do not succeed in optimizing safety investment benefits. Implementing precaution measures also goes hand in hand with costs. Nonetheless, to be able to take optimal prevention decisions, prevention costs should be weighed against hypothetical benefits resulting from the preventive measures. A cost-benefits analysis can be carried out to maximize an investment's efficiency, constrained by a given safety budget. To this end, prevention options should be intelligently compared with each other, on an aggregated level.

[3]Hypothetical benefits are benefits resulting from a nonarising accident and from the resulting consequential avoidance of related costs.

Table V. Quantifiable and Nonquantifiable Socioeconomic Consequences of Accidents

Interested Parties	Nonquantifiable Consequences of Accidents	Quantifiable Consequences of Accidents
Victim(s)	– Pain and suffering – Moral and psychic suffering – Loss of physical functioning – Loss of quality of life – Health and domestic problems – Reduced desire to work – Anxiety – Stress	– Loss of salary and bonuses – Limitation of professional skills – Time loss (medical treatment) – Financial loss – Extra costs
Colleagues	– Bad feeling – Anxiety or panic attacks – Reduced desire to work – Anxiety – Stress	– Time loss – Potential loss of bonuses – Heavier work load – Training and guidance of temporary employees
Organization	– Deterioration of social climate – Poor image	– Internal investigation – Damage to property and material – Reduction in productivity – Reduction in quality – New training for staff – Technical interference – Organizational costs – Higher production costs – Higher insurance premiums – Administrative costs – Sanctions imposed by parent company – Sanctions imposed by the government

Source: Based on De Greef and Van den Broek.[17]

3. APPROACH FOR COST-EFFICIENT PRECAUTION DECISIONS

3.1. Input Information

In industrial practice, companies are confronted with budget limitations. We call the available yearly budget for prevention related to safety Bu_{tot}. When possible prevention investments exceed this budget, they cannot be carried out. Therefore, only the preventative measures having a cost within Bu_{tot} will be considered in the approach.

To employ an approach to take cost-efficient prevention decisions, certain actions should have been carried out by the user and certain input information is needed. All risks should have been classified into one of the risk cells of the risk matrix. Every cell i corresponds to a potential cell cost C_i, determined by:

$$C_i = l_i \times c_i,$$

where C_i is costs resulting from an accident related to a risk from risk cell i, l_i is likelihood corresponding to risk cell i, and c_i is financial impact (consequences) corresponding to risk cell i.

Fig. 1 illustrates the cost figures (expressed in € per year) for the risk matrix (which is used in this article). Other risk matrix configurations are of course possible in real industrial practice. When precaution investments are made to decrease risks situated within cell i toward cell j (note that j is characterized with lower consequences and/or likelihood), the potential cell costs become C_j. Hypothetical benefits in that case can be calculated as $C_i - C_j$.

The required information for application of the approach is shown in Table VI.

When all these data are known, it is possible to use our approach to determine the most cost-efficient prevention measures. It should be noted that the information needed is at a risk-matrix-cell level, and as such, no individual risk information is needed, but only aggregated risk information. Hence, the scope of the suggested approach is on an aggregated scale and requires the user to estimate, in aggregated and composite terms, the prevention costs to go from one risk matrix cell to any other (lower) risk matrix cell

Table VI. Required Information for Application of the Approach

n
Number of cells where risks do exist for the organization ($= Nc; Nc \subset n$)
C_i
Bu_{tot}
Costs of prevention for going from risk cell i to risk cell j (CoP_{ij}), $\forall i, j$ whereby $i \in Nc$

(irrespective of the different kinds of risks that may be situated within the cells). The approach we elaborated is explained in the next section.

3.2. Approach Working Procedure and Illustrative Example

The first step in the first part of the approach is the categorization of (type I) risks into the risk classes of the risk matrix. We assume that Nc risk cells (out of the n risk cells in total) contain one or more risks. Prevention costs to go from risk cell i to risk cell j (note that $j < i$) are written as CoP_{ij}. If the prevention costs are higher than the yearly prevention budget Bu_{tot}, no investment will be made in these prevention measures; hence these prevention costs are excluded at the beginning of approach execution. The hypothetical benefits corresponding to a decrease in risk cell from i to j are calculated by subtracting C_j from C_i.

Following the analysis in the previous sections, a list of measures will have been drawn up. Using this list, the optimal risk portfolio can be determined using optimization. In its simplest form, determining the optimal risk portfolio is equal to solving a knapsack problem. The knapsack problem derives its name from the fact that a person having to fill his/her fixed-size knapsack with the most valuable items faces a similar problem. The knapsack problem is one of the most fundamental problems in combinatorial optimization and has many applications, for example, in stock portfolio management, as well as many extensions.

In the basic version of this problem, a set of decision variables x_i is defined where variable x_i (corresponding to measure i) takes on value 1 if this measure i is chosen as part of the portfolio and 0 if it is not. A mathematical formulation of the knapsack problem is the following:

$$\max B_i x_i$$

$$s.t. C_i x_i \leq Bu_{tot}$$

$$x_i \in \{0, 1\}$$

The first equation expresses the total benefit from the selected portfolio, which should be maximized. The second equation expresses the fact that the total cost of the selected measures should not exceed the budget. The third constraint implies that a measure is either fully taken or not taken at all.

A number of assumptions are implicitly taken in this formulation:

(1) A measure is either taken or not (it cannot be partially taken);
(2) The total benefit of all measures taken is the sum of the individual benefits of the chosen measures;
(3) The total cost of all measures taken is the sum of the costs of the individual measures;
(4) Measures can be independently implemented, without consequences for the other measures.

Some of these assumptions are not completely realistic. In the following section, we will discuss this.

Although the knapsack problem is NP hard,[4] it can be solved efficiently even for very large instances.[(24)] The advantage of using the knapsack-based formulation is that it can be solved by standard off-the-shelf commercial software for mixed-integer programming, such as CPLEX (http://www.ibm.com/software/integration/optimization/cplex-optimizer/) or Gurobi (http://www.gurobi.org) or their open source counterparts such as GLPK (http://www.gnu.org/software/glpk/) or lpsolve (http://lpsolve.sourceforge.net). Moreover, even spreadsheet software such as Excel or LibreOffice include a solver that can be used to approach and optimize the safety measures portfolio using the method described in this article.

Consider the following example to illustrate the approach. Note that the annual budgeting is assumed to be designed in such a way that no reservations can be made for capital-intensive items. On the basis of the risk matrix displayed in Fig. 1 and the basic

[4] An optimization problem is NP hard if the running time of the fastest known algorithm to solve it increases exponentially in the problem size. For example, if the problem size (the number of possible items in the knapsack) is n, the computing time required by the fastest known knapsack algorithm can be written as $a \cdot e^n$, where a is a constant.

Table VII. Information of Our Illustrative Example, for Application of the Approach

$n = 20$
$Nc = 6$: risk cells 3, 7, 10, 12, 13 and 15 from Fig. 2
C_i = see risk matrix from Fig. 1
$Bu_{tot} = €50{,}000$
Costs of prevention for going from risk cell *i* to risk cell *j* (CoP_{ij}) = see Table VIII

Table VIII. Costs of Prevention CoP_{ij} and Hypothetical Benefits for Our Illustrative Case

Prevention Measure *ij*	(Illustrative) Costs of Prevention for Going from *i* to *j* (CoP_{ij}) (€)	Hypothetical Benefits for Going from *i* to *j* (€)
Start = Risk cell 3		
3 2	35	67.5
3 1	42	74.25
Start = Risk cell 7		
7 6	325	675
7 5	460	742.5
7 3	295	675
7 2	420	742.5
7 1	590	749.25
Start = Risk cell 10		
10 9	330	675
10 6	350	675
10 5	390	742.5
10 2	400	742.5
10 1	880	749.25
Start = Risk cell 12		
12 11	13,500	17,500
12 10	13,750	24,250
12 9	14,800	24,925
12 8	13,000	22,500
12 7	15,000	24,250
12 6	16,500	24,925
12 5	26,000	24,992.5
12 4	13,900	24,750
12 3	17,000	24,925
12 2	27,500	24,992.5
12 1	38,000	24,999.25
Start = Risk cell 13		
13 9	410	675
13 5	550	742.5
13 1	700	749.25
Start = Risk cell 15		
15 14	31,000	67,500
15 13	36,650	74,250
15 11	29,880	67,500
15 10	38,000	74,250
15 9	52,000	74,925
15 7	41,440	74,250
15 6	48,990	74,925
15 5	64,450	74,992.5
15 3	50,000	74,925
15 2	62,250	74,992.5
15 1	88,000	74,999.25

information of Tables VII and VIII, the execution of the approach is explained.

An optimal allocation of safety measures with a maximum of one prevention measure from each of the risk cells, which can be assigned, needs to be determined. As explained earlier, to solve this problem, four conditions have to be met: (i) the total benefit of measures taken needs to be maximized; (ii) the available budget constraint needs to be respected; (iii) a maximum of one decrease per risk cell is allowed; and (iv) a measure can be taken or not. These conditions translate into the following mathematical expressions:

i) $\max_{i,j} \sum B_{ij} x_{ij}$
ii) $\sum_{i,j} CoP_{ij} \leq Bu_{tot}$
iii) $\sum_j x_{ij} \leq 1$
iv) $x_{ij}\ \{0, 1\}$

Solving the mathematical program in the aforementioned equations yields the optimal solution of the illustrative example represented by Tables VII and VIII. In this solution the measures taken are displayed in Fig. 3 with a total cost of €49,987, and a total hypothetical benefit of €97,499.25. The total hypothetical profit for the illustrative example thus equals €47,512.25.

It should be stressed that this illustrative example only serves the purpose of explaining the method of using the knapsack software to determine an optimal allocation of safety resources. It is possible and recommendable that exact figures are used to determine the hypothetical benefits, simply by calculating all the real costs per year of the risk cells for the feasible scenarios in the organization, and thus by creating a "real matrix" with real cell assignment figures (instead of general figures).

Furthermore, it is possible to further include more advanced conditions for real-case problems and situations within the method. The next section elaborates on what refinements might be carried out.

3.3. Approach Refinements for Further Application in Real Industrial Practice

3.3.1. Relationships Between Measures

In general, the portfolio of safety measures chosen by a company is subject to a number of extra constraints that express relationships between these

			chosen	cost	benefit
Start = Risk cell 3			1		
3 2	35	67.5	0	0	0
3 1	42	74.25	1	42	74.25
Start = Risk cell 7			1		
7 6	325	675	0	0	0
7 5	460	742.5	0	0	0
7 3	295	675	1	295	675
7 2	420	742.5	0	0	0
7 1	590	749.25	0	0	0
Start = Risk cell 10			0		
10 9	330	675	0	0	0
10 6	350	675	0	0	0
10 5	390	742.5	0	0	0
10 2	400	742.5	0	0	0
10 1	880	749.25	0	0	0
Start = Risk cell 12			1		
12 11	13500	17500	0	0	0
12 10	13750	24250	0	0	0
12 9	14800	24925	0	0	0
12 8	13000	22500	1	13000	22500
12 7	15000	24250	0	0	0
12 6	16500	24925	0	0	0
12 5	26000	24992.5	0	0	0
12 4	13900	24750	0	0	0
12 3	17000	24925	0	0	0
12 2	27500	24992.5	0	0	0
12 1	38000	24999.25	0	0	0
Start = Risk cell 13			0		
13 9	410	675	0	0	0
13 5	550	742.5	0	0	0
13 1	700	749.25	0	0	0
Start = Risk cell 15			1		
15 14	31000	67500	0	0	0
15 13	36650	74250	1	36650	74250
15 11	29880	67500	0	0	0
15 10	38000	74250	0	0	0
15 9	52000	74925	0	0	0
15 7	41440	74250	0	0	0
15 6	48990	74925	0	0	0
15 5	64450	74992.5	0	0	0
15 3	50000	74925	0	0	0
15 2	62250	74992.5	0	0	0
15 1	88000	74999.25	0	0	0
				49987	97499.25

Fig. 3. Solution of the illustrative example from Tables VII and VIII.

measures. Fortunately, these relationships are generally easily added to the knapsack-based approach, usually by introducing additional constraints. This section discusses some of these relationships and shows how they can be expressed in the approach using additional constraints.

3.3.2. Binary Relationships

If risk cell r is decreased, risk cell t also has to be decreased and vice versa. This situation occurs when measures are mutually dependent on each other and taking one measure without the other makes no sense. An example is when the use of a new device that enhances safety requires training. It does not make sense to install the device without the training, and it does not make sense to give the training without installing the device.

This relationship between risk cell decreases from r to s and from t to u can be expressed in the approach by the extra constraint:

$$x_{r\to s} = x_{t\to u}.$$

Another (less flexible) way to include this relationship in the approach is to combine risk cell decreases $r \to s$ and $t \to u$ in a single risk cell decrease.

Suppose, for example, that measures $15 \to 13$ and $13 \to 9$ either need to be taken together or not at all in our illustrative example. We add constraint $x_{15\to 13} = x_{13\to 9}$ to the knapsack problem and get solution $\{3 \to 1, 7 \to 1, 10 \to 1, 12 \to 9, 13 \to 1, 15 \to 14\}$ with total cost €48,012 and total hypothetical benefit €94,747. The total hypothetical profit in this case would be €44,735.

Another situation, that might occur is the following: if risk cell r is decreased, risk cell t also has to be decreased, but the reverse is not true. As an example, to prevent fire from spreading between departments, a company is considering installing a fire-resisting door. The time the door resists fire can be increased by adding an extra layer of fireproof coating to it. Clearly, applying the coating without installing the door makes no sense, but the reverse does.

The relationship between risk cell decrease $r \to s$ (installing the door) and risk cell decrease $t \to u$ (installing the fireproof coating) can be expressed as:

$$x_{r\to s} \le x_{t\to u}.$$

Suppose that when measure $15 \to 13$ is taken, measure $3 \to 2$ also needs to be taken. The constraint $x_{15\to 13} \le x_{3\to 2}$ is added, the problem is resolved, and the optimal solution is $\{3 \to 2, 7 \to 3, 12 \to 8, 15 \to 13\}$ with total cost €49,980 and total hypothetical benefit and profit, respectively, €97,492.5 and €47,512.5.

Yet another possible situation is that either risk cell r or risk cell t needs to be decreased, but not both risk cells at the same time. This situation can occur if two measures duplicate each other's effects and the company judges it superfluous to invest in both measures simultaneously. For example, a machine can be protected by a concrete casing or a steel casing, but not by both.

This can be mathematically expressed as follows:

$$x_{r\to s} = 1 - x_{t\to u}.$$

Another possibility is that either risk cell r, or risk cell t, or both, needs to be decreased. Such a situation can occur, for example, if the company management has decided that it will install at least one of smoke detectors or fire doors, but may also decide to install both. Translated into a mathematical constraint, this situation can be included in the safety measures allocation problem as follows:

$$x_{r\to s} + x_{t\to u} \ge 1.$$

Yet another feasible situation is that if risk cell t is decreased, risk cell r cannot be decreased and vice versa. But the possibility exists that both measures are not taken. This situation occurs, for example, when management has decided that smoke detectors might be installed, but two types are available and only one type will be selected at most.

This can be expressed as follows:

$$x_{r\to s} \le 1 - x_{t\to u}.$$

3.3.3. Other Relationships

In principle, all relationships between measures can be expressed as constraints in the knapsack problem. Essentially, the decision whether to decrease risk cell i can be seen as a literal in a propositional logic system, in which logical relationships are expressed by the operators NOT (risk cell i is not decreased), AND (risk cell i and risk cell j are decreased), OR (risk cell i or risk cell j is decreased), and IMPLICATION (if risk cell i is decreased, then risk cell j is decreased). These operators can be used to create arbitrarily complex relationships that can be used to express the most complex logical relationships between safety measures (e.g., if both the automatic fire door and the alarm system are installed, and the electricity system is not upgraded, then either a backup generator should be installed or a link to an additional power system should be purchased). Each of such relationships can be converted to constraints of the knapsack problem. We restrict ourselves to one example here. For a more elaborate discussion, including details on how to transform a logical relationship to constraints, we refer to Cavalier *et al.*,(25) Martello *et al.*,(24) Mendelson,(26) and Raman and Grossmann.(27)

Consider for example the following measures:

M_1: An automatic fire door is installed (e.g., $x_{4\to2}$).
M_2: An alarm system is installed (e.g., $x_{6\to2}$).
M_3: The electricity system is upgraded (e.g., $x_{3\to1}$).
M_4: A backup generator is installed (e.g., $x_{7\to3}$).
M_5: A link to an additional electricity system is installed (e.g., $x_{5\to2}$).

The condition that if both the automatic fire door and the alarm system are installed, and the electricity system is not upgraded, then either a backup generator should be installed or a link to an additional power system should be purchased, is logically equivalent to:[5]

$$(M_1 \wedge M_2) \wedge \neg M_3 \Rightarrow M_4 \vee M_5.$$

This can be converted into its conjunctive normal form:

$$(\neg M_1 \vee \neg M_2 \vee M_4 \vee M_5) \wedge (M_3 \vee M_4 \vee M_5),$$

which translates to the following two constraints:

$$\begin{cases} x_{3\to1} + x_{7\to3} + x_{5\to2} \geq 1 \\ x_{4\to2} + x_{6\to2} - x_{7\to3} - x_{5\to2} \leq 1 \end{cases}$$

3.3.4. *Nonadditivity*

For some situations, the benefits or costs of measures are not simply additive. Suppose, for example, that two fire doors can be installed in series to prevent fire from spreading to the next room. Clearly, the effect of installing one door instead of none will be larger than the effect of installing two doors instead of one. In other words, there will be a diminishing rate of return on the second door.

[5]The symbols in this equation are the following: ∧ is AND, ∨ is OR, ¬ is NOT, and ⇓ is IMPLIES.

This can be easily handled by identifying such situations and creating "virtual" measures in the cost-benefit table to represent the action of taking both measures. To ensure that each measure is only taken once, some additional constraints are also necessary.

As an example, suppose that the effect of combining risk cell decreases $3 \to 1$ and $7 \to 3$ in the example does not yield a benefit of $74.25 + 675 = 749.25$, but rather that it only yields 640. Suppose further that the cost of implementing both $3 \to 1$ and $7 \to 3$ is not $42 + 295 = 337$, but that a discount of €30 is given.

This can be handled by adding an extra risk cell decrease with cost 307 and hypothetical benefit 640. In addition, constraints are necessary to ensure that this extra measure is not taken if either $3 \to 1$ or $7 \to 3$ are taken. The additional constraints translate mathematically into:

$$x_{\text{extra risk cell decrease}} \leq 1 - x_{3\to1}$$
$$x_{\text{extra risk cell decrease}} \leq 1 - x_{7\to3}$$

In addition, we need to ensure that measures $3 \to 1$ and $7 \to 3$ are not both chosen at the same time:

$$x_{3\to1} + x_{7\to3} \leq 1.$$

In such case, resolving our illustrative example leads to an optimal solution, where the following risk reductions are carried out: $3 \to 2$, $12 \to 8$, and $15 \to 13$. The solution displays a total cost equal to €49,992 and a total hypothetical benefit of €97,817.5 and a total hypothetical profit of €47,825.5.

4. CONCLUSIONS

Preventing accidents that tend to happen relatively regularly in the industry is an important expenditure on a yearly basis. Optimizing prevention investments and making investment decisions in a cost-efficient way is therefore essential for corporations. To this end, we suggest a user-friendly knapsack-based approach to take cost-efficient prevention decisions. The approach employs some essential data that can easily be determined by any organization and that can be displayed using a risk matrix. Prevention costs are weighed against hypothetical benefits following the preventive measures taken, and the most cost-efficient preventive measures are determined following the knapsack algorithm, given a certain prevention budget available.

REFERENCES

1. Greenberg HR, Cramer JJ. Risk Assessment and Risk Management for the Chemical Process Industry. Toronto: John Wiley and Sons, 1991.
2. CCPS (Center for Chemical Process Safety). Guidelines for Hazard Evaluation Procedures, 2nd ed. New York: American Institute of Chemical Engineers, 1992.
3. Casal J. Evaluation of the Effects and Consequences of Major Accidents in Industrial Plants. Amsterdam: Elsevier, 2008.
4. Taleb NN. The Black Swan: The Impact of the Highly Improbable. New York: Random House, 2007.
5. Cox LA, Jr. Evaluating and improving risk formulas for allocating limited budgets to expensive risk-reduction opportunities. Risk Analysis, 2012; 32(7):1244–1252.
6. Cook R. Simplifying the creation and use of the risk matrix. Pp. 239–263 in Redmill F, Anderson T (eds). Improvements in System Safety. Bristol: Springer, 2007.
7. Wilkinson G, David R. Back to the basics: Risk matrices and ALARP. Pp. 179–182 in Dale C, Anderson T (eds). Safety-Critical System: Problems, Process and Practice. London: Springer, 2009.
8. Cox LA, Jr. What's wrong with risk matrices. Risk Analysis, 2008; 28(2):497–512.
9. Smith E, Siefert W, Drain D. Risk matrix input data biases. Systems Engineering, 2009; 12(4):344–360.
10. Middleton M, Franks A. Using risk matrices. Chemical Engineer, 2001; 723:34–37.
11. Department of Defense. Standard Practice for System Safety. MIL-STD-882D. Ohio: HQ Air Force Materiel Command, 2000.
12. Bridges WG, Williams TR. Risk acceptance criteria and risk judgment tools applied worldwide within a chemical company. In Proceedings of the International Conference and Workshop on Risk Analysis in Process Safety. Atlanta: American Institute of Chemical Engineers, 1997.
13. Gillett J. Rapid ranking of hazards. Process Engineering, 1985; 66(2):19–22.
14. Moses KDPE, Malone RW, Jr. Development of risk assessment matrix for NASA engineering and safety center. Pp. 1–22 in Sage AP, Rouse WP (eds). Handbook of System Engineering and Management. New York: Wiley, 2004.
15. Levner E, Ganoulis J, Linkov I, Benayahu Y. Multiobjective risk/cost analysis of artificial marine systems using decision trees and fuzzy expert estimations. Pp. 161–174 in Linkov I, Kiker GA, Wenning RJ (eds). Environmental Security in Harbors and Coastal Areas. Heidelberg: Springer, 2007.
16. Reniers GLL, Soudan K. Risicoanalyse procedures in de scheikundige nijverheid: resultaten van kwalitatief onderzoek bij 24 chemische plants. Economisch en Sociaal Tijdschrift, 2003; 57(3):249–274.
17. De Greef M, Van den Broeck K. Report: Making the Case for Workplace Health Promotion. Essen: European Network for Workplace Health Promotion, 2004.
18. Hanley N, Spash CL. Cost-Benefit Analysis and the Environment. Cheltenham: Edward Elgar Publishing, 1993.
19. Mossink J, Licher F (eds). Costs and Benefits of Occupational Safety and Health. In Proceedings of the European Conference on Costs and Benefits of Occupational Safety and Health 1997. The Hague, The Netherlands, 1997.
20. Rikhardsson PM, Impgaard M. Corporate cost of occupational accidents: An activity-based analysis. Accident Analysis and Prevention, 2004; 36(2):173–182.
21. Leigh JP, Waehrer G, Miller TR, Keenan C. Costs of occupational injury and illnesses across industries. Scandinavian Journal of Work, Environment and Health, 2003; 30(3):199–205.
22. Reniers GLL, Audenaert A. Chemical plant innovative safety investments decision-support methodology. Journal of Safety Research, 2009; 40(6):411–419.
23. Caputo AC, Palumbo M, Pelagagge PM. Economic optimization of industrial safety measures using genetic algorithms. Journal of Loss Prevention in the Process Industries, 2011; 24(5):541–551.
24. Martello S, Pisinger D, Toth P. New trends in exact algorithms for the 0–1 knapsack problem. European Journal of Operational Research, 2000; 123(2):325–332.
25. Cavalier TM, Pardalos PM, Soyster AL. Approaching and integer programming techniques applied to propositional calculus. Computers and Operations Research, 1990; 17(6):561–570.
26. Mendelson E. Introduction to Mathematical Logic, 4th ed. London: Chapman & Hall, 1997.
27. Raman R, Grossmann IE. Relation between milp modeling and logical inference for chemical process synthesis. Computers and Chemical Engineering, 1991; 15(2):73–84.